



A Robbins-Monro procedure for a class of models of deformation

Philippe Fraysse

► To cite this version:

Philippe Fraysse. A Robbins-Monro procedure for a class of models of deformation. 2012. hal-00696050

HAL Id: hal-00696050

<https://hal.science/hal-00696050>

Preprint submitted on 10 May 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A ROBBINS-MONRO PROCEDURE FOR A CLASS OF MODELS OF DEFORMATION

PHILIPPE FRAYSSE

Université de Bordeaux

ABSTRACT. The paper deals with the statistical analysis of several data sets associated with shape invariant models with different translation, height and scaling parameters. We propose to estimate these parameters together with the common shape function. Our approach extends the recent work of Bercu and Fraysse to multivariate shape invariant models. We propose a very efficient Robbins-Monro procedure for the estimation of the translation parameters and we use these estimates in order to evaluate scale parameters. The main pattern is estimated by a weighted Nadaraya-Watson estimator. We provide almost sure convergence and asymptotic normality for all estimators. Finally, we illustrate the convergence of our estimation procedure on simulated data as well as on real ECG data.

1. INTRODUCTION

Statistic analysis of models with periodic data is a mathematical field of great interest. Indeed, a detailed analysis of such models enables us to have a satisfactory approximation of real life phenomena. In particular, semiparametric models are often used to describe a large number of phenomena as Meteorology [22], road traffic [2], [7] or children growth [8]. More precisely, all these examples are modeled by a particular subclass of semiparametric models, called shape invariant models [13]. Periodic shape invariant model is a semiparametric regression model with an unknown periodic shape function. In this paper, we consider several data sets associated with a common shape function and differing from each other by three parameters, a translation, a height and a scale. Formally, we are interested in the following shape invariant model

$$(1.1) \quad Y_{i,j} = a_j f(X_i - \theta_j) + v_j + \varepsilon_{i,j},$$

where $1 \leq i \leq n$ and $1 \leq j \leq p$ and the common shape function f is periodic. The variables X_i are random, which generalizes the case of equi-distributed deterministic variables studied for example in [7], [20] or [22]. The case where $p = 1$, $a_1 = 1$ and $v_1 = 0$ has been previously investigated by Bercu and Fraysse [1] and we propose to generalize their study to the multivariate case $p \geq 2$ and nontrivial scale and height parameters. We are interested in the parametric estimation of the height parameters

2010 *Mathematics Subject Classification.* Primary: 62G05, Secondary: 62G20.

Key words and phrases. Semiparametric estimation, estimation of shifts, estimation of a regression function, asymptotic properties.

v , the translation parameters θ and the scale parameters a , respectively given by

$$(1.2) \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_p \end{pmatrix}, \quad \theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_p \end{pmatrix}, \quad a = \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix}.$$

Our goal is also to estimate the common shape function f . Estimation of shifts has lots of similarities with curve registration and alignment problems [18]. Analysis of ECG curves [20] or the study of traffic data [2], [7] fall into this framework. In [8], Gasser and Kneip propose to estimate the shifts by aligning the maxima of the curves, their position being estimated by the zeros of the derivative of a kernel estimate. In the case where $a_j = 1$ and $v_j = 0$, Gamboa *et al.* [7] provide a semiparametric method for the estimation of the shifts. They use a Discrete Fourier Transform to transform the model (1.1) into the Fourier domain. Then, they estimate the shift parameters from the transformed model using M -estimation. This M -estimate converges almost surely as the number of Fourier coefficients increases, and they establish asymptotic normality of their estimator. The important contribution of Vimond [22] generalizes this study, adding the estimation of scale and height parameters. When the parameters θ_j are supposed to be random, Castillo and Loubes [2] provide sharp estimators of the θ_j , following the approach of Dalalyan *et al.* [4] in the Gaussian white noise framework. Then, they recover the unknown density of the θ_j using a kernel density estimator. Trigano *et al.* [20] follow the same approach and provide an application on ECG data. Note that they prove asymptotic results when p goes to infinity contrary to [7] whose asymptotic results are established when n goes to infinity. In this paper, we shall focus our attention on asymptotic results related to n . More precisely, for the estimation of the θ_j 's, we propose to make use of a multidimensional version of the Robbins-Monro algorithm [19]. Assume that one can find a function $\phi : \mathbb{R}^p \rightarrow \mathbb{R}^p$, free of the parameter θ , such that $\phi(\theta) = 0$. Then, it is possible to estimate θ by the Robbins-Monro algorithm

$$(1.3) \quad \hat{\theta}_{n+1} = \hat{\theta}_n + \gamma_n T_{n+1}$$

where (γ_n) is a positive sequence of real numbers decreasing towards zero and (T_n) is a sequence of random vectors such that $\mathbb{E}[T_{n+1} | \mathcal{F}_n] = \phi(\hat{\theta}_n)$ where \mathcal{F}_n stands for the σ -algebra of the events occurring up to time n . Under standard conditions on the function ϕ and on the sequence (γ_n) , it is well-known [6] that $\hat{\theta}_n$ tends to θ almost surely. The asymptotic normality of $\hat{\theta}_n$ may be found in [16] whereas the quadratic strong law and the law of iterated logarithm are established in [15]. Results for randomly truncated version of the Robbins-Monro algorithm are given in [12].

The second part of the paper concerns the nonparametric estimation of the regression function f which represents the common behavior of every subject j . A wide range of literature is available on nonparametric estimation of a regression function. We refer the reader to [5], [21] for two excellent books on density and regression function estimation. Here, we focus our attention on the Nadaraya-Watson estimator [14] [24] of f . Our approach follows essentially the same lines as in [1].

More precisely, we propose to make use of a weighted recursive Nadaraya-Watson estimator [6] of f which takes into account the previous estimation of v , θ and a , respectively by \widehat{v}_n , $\widehat{\theta}_n$ and \widehat{a}_n . It is given, for all $x \in \mathbb{R}$, by

$$(1.4) \quad \widehat{f}_n(x) = \sum_{j=1}^p \omega_j(x) \widehat{f}_{n,j}(x)$$

with

$$\widehat{f}_{n,j}(x) = \frac{1}{\widehat{a}_{n,j}} \frac{\sum_{i=1}^n W_{i,j}(x) (Y_{i,j} - \widehat{v}_{i-1,j})}{\sum_{i=1}^n W_{i,j}(x)}$$

and

$$W_{n,j}(x) = \frac{1}{h_n} K\left(\frac{X_n - \widehat{\theta}_{n-1,j} - x}{h_n}\right),$$

where $\widehat{v}_{n,j}$, $\widehat{\theta}_{n-1,j}$ and $\widehat{a}_{n,j}$ are respectively the j -th component of \widehat{v}_n , $\widehat{\theta}_{n-1}$ and \widehat{a}_n . Moreover, the kernel K is a chosen probability density function and the bandwidth (h_n) is a sequence of positive real numbers decreasing to zero. The main difficulty arising here is that we have to deal with the additional term $\widehat{\theta}_n$ inside the kernel K .

The paper is organized as follows. Section 2 presents the model and the hypothesis which are necessary to carry out our statistical analysis. Section 3 is devoted to the parametric estimation of the vector v , while Section 4 deals with our Robbins-Monro procedure for the parametric estimation of θ . Section 5 concerns the parametric estimation of the vector a . In these three sections, we establish the almost sure convergence of \widehat{v}_n , $\widehat{\theta}_n$ and \widehat{a}_n as well as their asymptotic normality. A quadratic strong law is also provided for these three estimates. Section 6 deals with the nonparametric estimation of f . Under standard regularity assumptions on the kernel K , we prove the almost sure pointwise convergence of \widehat{f}_n to f . In addition, we also establish the asymptotic normality of \widehat{f}_n . Section 7 contains some numerical experiments on simulated and real data, illustrating the performances of our semiparametric estimation procedure. The proofs of the parametric results are given in Section 8, while those concerning the nonparametric results are postponed to Section 9. Finally, Section 10 is devoted to the identifiability constraints associated with the model (1.1).

2. MODEL AND HYPOTHESIS

We shall propose several natural identifiability constraints associated with the model (1.1) given, for all $1 \leq i \leq n$ and $1 \leq j \leq p$, by

$$Y_{i,j} = a_j f(X_i - \theta_j) + v_j + \varepsilon_{i,j},$$

where, for all $1 \leq j \leq p$, $a_j \neq 0$. The noise $(\varepsilon_{i,j})$ is a sequence of independent random variables with mean zero and variances $\mathbb{E}[\varepsilon_{i,j}^2] = \sigma_j^2 < +\infty$. In addition, as in [1], we make the following hypothesis.

- (\mathcal{H}_1) The observation times (X_i) are independent and identically distributed with probability density function g , positive on its support $[-1/2; 1/2]$. In addition, g is continuous, twice differentiable with bounded derivatives.
- (\mathcal{H}_2) The shape function f is symmetric, bounded, periodic with period 1.

Our goal is to estimate the parameters a, θ, v as well as the common shape function f . However, the shape invariant model (1.1) is not always identifiable. Indeed, for a given vector of parameters (a, θ, v) and a given shape function f , one can find another vector of parameters (a^*, θ^*, v^*) and an other shape function f^* such that for all $1 \leq j \leq p$ and for all $x \in \mathbb{R}$,

$$(2.1) \quad a_j f(x - \theta_j) + v_j = a_j^* f^*(x - \theta_j^*) + v_j^*.$$

Härdle and Marron [10] or Kneip and Engel [11] were among the first to discuss about identifiability. However, in most papers dealing with shape invariant models, the identifiability is not so clear. Indeed, the identifiability conditions proposed by some authors could be not as restrictive as it should be. For example, under the conditions provided by Vimond [22], it is possible to find two different vectors of parameters (a, θ, v) and (a^*, θ^*, v^*) satisfying their conditions and (2.1), while under the conditions provided by Wang *et al.* [23], it is possible to find two different shape functions f and f^* satisfying their conditions and (2.1). That is the reason why we have chosen carefully the identifiability constraints. More precisely, we impose the following conditions.

$$\begin{aligned} (\mathcal{H}_3) \quad & \int_{-1/2}^{1/2} f(x) dx = 0, \\ (\mathcal{H}_4) \quad & a_1 = 1, \theta_1 = 0 \text{ and } \max_{1 \leq j \leq p} |\theta_j| < 1/4. \end{aligned}$$

Hypothesis (\mathcal{H}_3) allows us to define uniquely the v_j and the second constraint in (\mathcal{H}_4) define uniquely the θ_j , while $a_1 = 1, \theta_1 = 0$ implies that the first curve is taken as a reference. These conditions are well-adapted to our framework. Note that (\mathcal{H}_3) and (\mathcal{H}_4) could be replaced by (\mathcal{H}'_3) and (\mathcal{H}'_4) defined as

$$\begin{aligned} (\mathcal{H}'_3) \quad & \int_{-1/2}^{1/2} f(x) dx = 0 \text{ and } \sup_{x \in [0;1]} |f(x)| = 1, \\ (\mathcal{H}'_4) \quad & \theta_1 = 0, \min_{1 \leq j \leq p} a_j > 0 \text{ and } \max_{1 \leq j \leq p} |\theta_j| < 1/2. \end{aligned}$$

An other alternative could be to substitute the hypothesis (\mathcal{H}_4) by (\mathcal{H}''_4) or (\mathcal{H}'''_4) defined as

$$\begin{aligned} (\mathcal{H}''_4) \quad & a_1 = 1, \theta_1 = 0, \min_{1 \leq j \leq p} a_j > 0, \max_{1 \leq j \leq p} |\theta_j| < 1/2, \\ (\mathcal{H}'''_4) \quad & a_1 > 0, \theta_1 = 0, \sum_{j=1}^p a_j^2 = p \text{ and } \max_{1 \leq j \leq p} |\theta_j| < 1/4. \end{aligned}$$

Finally, note that the hypothesis of symmetry of f is not necessary to ensure identifiability of the model (1.1) and it is possible to get rid of the symmetry assumption on f , see Remark 4.3 below. Nevertheless, this hypothesis makes the estimation of the internal shifts easier, as we shall see in Section 4. All those identifiability conditions are discussed in Section 10.

3. ESTIMATION OF THE EXTERNAL SHIFTS

Via (\mathcal{H}_2) and (\mathcal{H}_3) , it is not difficult to see that

$$\mathbb{E} \left[\frac{Y_{i,j}}{g(X_i)} \right] = v_j.$$

Then, a natural consistent estimator \hat{v}_n of v is given, for all $1 \leq j \leq p$, by

$$(3.1) \quad \hat{v}_{n,j} = \frac{1}{n} \sum_{i=1}^n \frac{Y_{i,j}}{g(X_i)}.$$

In order to establish the asymptotic behaviour of \hat{v}_n , we denote by Y the vector

$$(3.2) \quad Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_p \end{pmatrix},$$

where

$$Y_j = a_j f(X - \theta_j) + v_j + \varepsilon_j,$$

with X a random variable sharing the same distribution as the sequence (X_n) and for $1 \leq j \leq p$, ε_j sharing the same distribution as the sequence $(\varepsilon_{i,j})$.

The asymptotic results for \hat{v}_n are as follows.

Theorem 3.1. *Assume that (\mathcal{H}_1) to (\mathcal{H}_4) hold. Then, we have the a.s. convergence*

$$(3.3) \quad \lim_{n \rightarrow +\infty} \hat{v}_n = v \quad \text{a.s.}$$

and the asymptotic normality

$$(3.4) \quad \sqrt{n} (\hat{v}_n - v) \xrightarrow{\mathcal{L}} \mathcal{N}_p(0, \Gamma(v)),$$

where $\Gamma(v)$ stands for the covariance matrix given by

$$\Gamma(v) = \text{Cov} \left(\frac{Y}{g(X)} \right).$$

In addition, we also have the quadratic strong law

$$(3.5) \quad \lim_{n \rightarrow +\infty} \frac{1}{\log(n)} \sum_{i=1}^n (\hat{v}_i - v) (\hat{v}_i - v)^T = \Gamma(v) \quad \text{a.s.}$$

4. ESTIMATION OF THE INTERNAL SHIFTS

In all the sequel, we introduce an auxiliary function ϕ defined for all $t \in \mathbb{R}^p$, by

$$(4.1) \quad \phi(t) = \mathbb{E} \left[D(X, t) \begin{pmatrix} a_1 f(X - \theta_1) \\ \vdots \\ a_p f(X - \theta_p) \end{pmatrix} \right]$$

where $D(X, t)$ stands for the diagonal square matrix of order p defined by

$$(4.2) \quad D(X, t) = \frac{1}{g(X)} \text{diag} \left(\sin(2\pi(X - t_1)), \dots, \sin(2\pi(X - t_p)) \right).$$

Using the symmetry of f , the same calculations as in [1] lead, for all $1 \leq j \leq p$, to

$$(4.3) \quad \mathbb{E} \left[\frac{\sin(2\pi(X - t_j))}{g(X)} a_j f(X - \theta_j) \right] = a_j f_1 \sin(2\pi(\theta_j - t_j))$$

where f_1 stands for the first Fourier coefficient of f

$$f_1 = \int_{-1/2}^{1/2} \cos(2\pi x) f(x) dx.$$

Consequently,

$$(4.4) \quad \phi(t) = f_1 \begin{pmatrix} a_1 \sin(2\pi(\theta_1 - t_1)) \\ \vdots \\ a_p \sin(2\pi(\theta_p - t_p)) \end{pmatrix}.$$

Throughout the paper, in order to avoid tedious discussion, we assume that $f_1 \neq 0$. Roughly speaking, we are going to implement our Robbins-Monro procedure as in [1] for each component of θ . More precisely, for all $1 \leq j \leq p$, by noting $\phi_j(t) = a_j f_1 \sin(2\pi(\theta_j - t_j))$, if $|t_j - \theta_j| < 1/2$, $(t_j - \theta_j)\phi_j(t) < 0$ if $\text{sign}(a_j f_1) > 0$ and $(t_j - \theta_j)\phi_j(t) > 0$ otherwise. Moreover, denote $K = [-1/4; 1/4]$. Then, we define the projection of $x \in \mathbb{R}$ on K by

$$\pi_K(x) = \begin{cases} x & \text{if } |x| \leq 1/4, \\ 1/4 & \text{if } x \geq 1/4, \\ -1/4 & \text{if } x \leq -1/4. \end{cases}$$

Let (γ_n) be a decreasing sequence of positive real numbers satisfying

$$(4.5) \quad \sum_{n=1}^{\infty} \gamma_n = +\infty \quad \text{and} \quad \sum_{n=1}^{\infty} \gamma_n^2 < +\infty.$$

For the sake of clarity, we shall make use of $\gamma_n = 1/n$. Then, for $1 \leq j \leq p$, we estimate θ_j via the sequence $(\hat{\theta}_{n,j})$ defined, for all $n \geq 1$, by

$$(4.6) \quad \hat{\theta}_{n+1,j} = \pi_K \left(\hat{\theta}_{n,j} + \gamma_{n+1} \text{sign}(a_j f_1) T_{n+1,j} \right)$$

where the initial value $\hat{\theta}_0 \in K^p$ and the random vector T_{n+1} is given by

$$(4.7) \quad T_{n+1} = D(X_{n+1}, \hat{\theta}_n) \begin{pmatrix} Y_{n+1,1} \\ \vdots \\ Y_{n+1,p} \end{pmatrix}.$$

The almost sure convergence for the estimator $\hat{\theta}_n$ is as follows.

Theorem 4.1. *Assume that (\mathcal{H}_1) to (\mathcal{H}_4) hold. Then, $\widehat{\theta}_n$ converges almost surely to θ as n goes to $+\infty$. In addition, for $1 \leq j \leq p$, the number of times that the random variable $\widehat{\theta}_{n,j} + \gamma_{n+1} \text{sign}(a_j f_1) T_{n+1,j}$ goes outside of K is almost surely finite.*

Remark 4.1. *At first sight, the estimation procedure needs the knowledge of the sign of $a_j f_1$ for all $1 \leq j \leq p$. However, it is possible to do without. Indeed, denote by $(\widehat{\theta}_{n,j}^+)$ the sequence defined, for $n \geq 1$, as*

$$\widehat{\theta}_{n+1,j}^+ = \pi_K \left(\widehat{\theta}_{n,j}^+ + \gamma_{n+1} T_{n+1,j}^+ \right),$$

where

$$T_{n+1}^+ = D(X_{n+1}, \widehat{\theta}_n^+) \begin{pmatrix} Y_{n+1,1} \\ \vdots \\ Y_{n+1,p} \end{pmatrix},$$

and by $\widehat{\theta}_{n,j}^-$ the sequence defined, for $n \geq 1$, as

$$\widehat{\theta}_{n+1,j}^- = \pi_K \left(\widehat{\theta}_{n,j}^- - \gamma_{n+1} T_{n+1,j}^- \right),$$

where

$$T_{n+1}^- = D(X_{n+1}, \widehat{\theta}_n^-) \begin{pmatrix} Y_{n+1,1} \\ \vdots \\ Y_{n+1,p} \end{pmatrix}.$$

Then, two events are possible. More precisely, for $1 \leq j \leq p$,

$$\lim_{n \rightarrow +\infty} \widehat{\theta}_{n,j}^+ = \theta_j \quad \text{and} \quad \lim_{n \rightarrow +\infty} |\widehat{\theta}_{n,j}^-| = 1/4 \quad \text{a.s.}$$

or

$$\lim_{n \rightarrow +\infty} \widehat{\theta}_{n,j}^- = \theta_j \quad \text{and} \quad \lim_{n \rightarrow +\infty} |\widehat{\theta}_{n,j}^+| = 1/4 \quad \text{a.s.}$$

Hence, for n large enough, the vector $\widehat{\theta}_n$ which is considered is the vector whose absolute value of the j -th component is given by $\min(|\widehat{\theta}_{n,j}^+|, |\widehat{\theta}_{n,j}^-|)$. Nevertheless, for the sake of clarity, we shall do as if the sign of $a_j f_1$ is known.

In order to establish the asymptotic normality of $\widehat{\theta}_n$, it is necessary to introduce a second auxiliary function φ defined, for all $t \in \mathbb{R}^p$, as

$$(4.8) \quad \varphi(t) = \mathbb{E}[V(t)V(t)^T]$$

where $V(t)$ is given by

$$V(t) = \text{diag}(\text{sign}(a_1 f_1), \dots, \text{sign}(a_p f_1)) D(X, t) Y,$$

with Y given by (3.2). As soon as $4\pi|f_1| \min_{1 \leq j \leq p} |a_j| > 1$, denote for all $1 \leq k, l \leq p$,

$$\Sigma(\theta)_{k,l} = \frac{\varphi(\theta)_{k,l}}{2\pi(|a_k| + |a_l|)|f_1| - 1}.$$

Theorem 4.2. Assume that (\mathcal{H}_1) to (\mathcal{H}_4) hold. In addition, suppose that $(\varepsilon_{i,j})$ has a finite moment of order > 2 and that

$$4\pi|f_1|\min_{1 \leq j \leq p}|a_j| > 1.$$

Then, we have the asymptotic normality

$$(4.9) \quad \sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}_p(0, \Sigma(\theta)).$$

Theorem 4.3. Assume that (\mathcal{H}_1) to (\mathcal{H}_4) hold. In addition, suppose that $(\varepsilon_{i,j})$ has a finite moment of order > 2 and that

$$4\pi|f_1|\min_{1 \leq j \leq p}|a_j| > 1.$$

Then, we have the law of iterated logarithm, given, for all $w \in \mathbb{R}^p$, by

$$(4.10) \quad \begin{aligned} \limsup_{n \rightarrow \infty} \left(\frac{n}{2 \log \log n} \right)^{1/2} w^T (\hat{\theta}_n - \theta) &= - \liminf_{n \rightarrow \infty} \left(\frac{n}{2 \log \log n} \right)^{1/2} w^T (\hat{\theta}_n - \theta) \\ &= \sqrt{w^T \Sigma(\theta) w} \quad \text{a.s.} \end{aligned}$$

In particular,

$$(4.11) \quad \limsup_{n \rightarrow \infty} \left(\frac{n}{2 \log \log n} \right) (\hat{\theta}_n - \theta) (\hat{\theta}_n - \theta)^T = \Sigma(\theta) \quad \text{a.s.}$$

In addition, we also have the quadratic strong law

$$(4.12) \quad \lim_{n \rightarrow \infty} \frac{1}{\log n} \sum_{i=1}^n (\hat{\theta}_i - \theta)(\hat{\theta}_i - \theta)^T = \Sigma(\theta) \quad \text{a.s.}$$

Remark 4.2. In the particular case where

$$4\pi|f_1|\min_{1 \leq j \leq p}|a_j| = 1,$$

it is also possible to show, thanks to Theorem 2.2.12 page 52 of [6], that

$$\sqrt{\frac{n}{\log(n)}}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \varphi(\theta)).$$

Remark 4.3. As in [1], it is also possible to get rid of the symmetry assumption on f . However, it requires the knowledge of the first Fourier coefficients of f

$$f_1 = \int_{-1/2}^{1/2} \cos(2\pi x) f(x) dx \quad \text{and} \quad g_1 = \int_{-1/2}^{1/2} \sin(2\pi x) f(x) dx.$$

On the one hand, it is necessary to assume that $f_1 \neq 0$ or $g_1 \neq 0$, and to replace in (4.1) the diagonal matrix $D(X, t)$ defined in (4.2), by

$$(4.13) \quad \Delta(X, t) = \frac{1}{g(X)} \text{diag}(\delta(X, t_1), \dots, \delta(X, t_p)),$$

where, for all $1 \leq j \leq p$,

$$\delta(X, t_j) = f_1 \sin(2\pi(X - t_j)) - g_1 \cos(2\pi(X - t_j)).$$

Then, Theorem 4.1 is true for the projected Robbins-Monro algorithm defined, for all $1 \leq j \leq p$, by

$$\hat{\theta}_{n+1,j} = \pi_K \left(\hat{\theta}_{n,j} + \text{sign}(a_j) \gamma_{n+1} T_{n+1,j} \right),$$

where the initial value $\hat{\theta}_0 \in K^p$ and the random vector T_{n+1} is given by

$$T_{n+1} = \Delta(X_{n+1}, \hat{\theta}_n) \begin{pmatrix} Y_{n+1,1} \\ \vdots \\ Y_{n+1,p} \end{pmatrix}.$$

On the other hand, we also have to replace the second function φ defined in (4.8) by

$$\Psi(t) = \mathbb{E}[W(t)W(t)^T],$$

where $W(t)$ is given by

$$W(t) = \text{diag}(\text{sign}(a_1 f_1), \dots, \text{sign}(a_p f_1)) \Delta(X, t) Y$$

Then, as soon as $4\pi(f_1^2 + g_1^2) \min_{1 \leq j \leq p} |a_j| > 1$, Theorems 4.2 and 4.3 hold with

$$\Sigma(\theta)_{k,l} = \frac{\Psi(\theta)_{k,l}}{2\pi(f_1^2 + g_1^2)(|a_k| + |a_l|) - 1}.$$

In the rest of the paper, we shall not go in that direction as our strategy is to make very few assumptions on the Fourier coefficients of f .

5. ESTIMATION OF THE SCALE PARAMETERS

Henceforth, we introduce an other auxiliary function ψ defined for all $t \in \mathbb{R}^p$, by

$$(5.1) \quad \psi(t) = \mathbb{E} \left[C(X, t) \begin{pmatrix} a_1 f(X - \theta_1) \\ \vdots \\ a_p f(X - \theta_p) \end{pmatrix} \right]$$

where $C(X, t)$ is the diagonal matrix of order p , given by

$$(5.2) \quad C(X, t) = \frac{1}{g(X)} \text{diag}(\cos(2\pi(X - t_1)), \dots, \cos(2\pi(X - t_p))).$$

As for (4.4), we have

$$(5.3) \quad \psi(t) = f_1 \begin{pmatrix} a_1 \cos(2\pi(\theta_1 - t_1)) \\ \vdots \\ a_p \cos(2\pi(\theta_p - t_p)) \end{pmatrix}.$$

Then, it is clear from Theorem 4.1 that $\psi(\hat{\theta}_n)$ tends to $\psi(\theta) = f_1 a$. Hence, denote by (\hat{a}_n) the sequence defined, for $n \geq 1$ and for all $1 \leq j \leq p$, as

$$(5.4) \quad \hat{a}_{n,j} = \frac{1}{nf_1} \sum_{i=1}^n \frac{\cos(2\pi(X_i - \hat{\theta}_{i-1,j}))}{g(X_i)} Y_{i,j}.$$

The asymptotic behaviour of \hat{a}_n is as follows.

Theorem 5.1. *Assume that (\mathcal{H}_1) to (\mathcal{H}_4) hold. Then, we have the a.s. convergence*

$$(5.5) \quad \lim_{n \rightarrow +\infty} \widehat{a}_n = a \quad \text{a.s.}$$

and the asymptotic normality

$$(5.6) \quad \sqrt{n} (\widehat{a}_n - a) \xrightarrow{\mathcal{L}} \mathcal{N}_p(0, \Gamma(a)),$$

where $\Gamma(a)$ stands for the covariance matrix given by

$$(5.7) \quad \Gamma(a) = \frac{1}{f_1^2} \text{Cov}(C(X, \theta)Y).$$

In addition, we also have the quadratic strong law

$$(5.8) \quad \lim_{n \rightarrow +\infty} \frac{1}{\log(n)} \sum_{i=1}^n (\widehat{a}_i - a) (\widehat{a}_i - a)^T = \Gamma(a) \quad \text{a.s.}$$

Remark 5.1. *The estimation procedure needs the estimation of the first Fourier coefficient f_1 . Under the hypothesis (\mathcal{H}_4) , it is natural to estimate f_1 by*

$$\widehat{f}_{1,n} = \frac{1}{n} \sum_{i=1}^n \frac{\cos(2\pi X_i)}{g(X_i)} Y_{i,1}.$$

6. ESTIMATION OF THE REGRESSION FUNCTION

In this section, we are interested in the nonparametric estimation of the regression function f via a recursive Nadaraya-Watson estimator. On the one hand, we add the following standard hypothesis.

(\mathcal{H}_5) The regression function f is Lipschitz.

On the other hand, we follow the same approach as in [1]. Moreover, for more accuracy, we consider a weighted version of the Nadaraya-Watson estimator

$$(6.1) \quad \widehat{f}_n(x) = \sum_{j=1}^p \omega_j(x) \widehat{f}_{n,j}(x),$$

where, for all $1 \leq j \leq p$,

$$(6.2) \quad \omega_j(x) \geq 0 \quad \text{and} \quad \sum_{j=1}^p \omega_j(x) = 1,$$

$$(6.3) \quad \widehat{f}_{n,j}(x) = \frac{1}{\widehat{a}_{n,j}} \frac{\sum_{i=1}^n (W_{i,j}(x) + W_{i,j}(-x)) (Y_{i,j} - \widehat{v}_{i-1,j})}{\sum_{i=1}^n (W_{i,j}(x) + W_{i,j}(-x))},$$

and

$$(6.4) \quad W_{n,j}(x) = \frac{1}{h_n} K\left(\frac{X_n - \widehat{\theta}_{n-1,j} - x}{h_n}\right).$$

The bandwidth (h_n) is a sequence of positive real numbers, decreasing to zero, such that nh_n tends to infinity. For the sake of simplicity, we propose to make use of

$h_n = 1/n^\alpha$ with $\alpha \in]0, 1[$. Moreover, we shall assume in all the sequel that the kernel K is a positive symmetric function, bounded with compact support, satisfying

$$\int_{\mathbb{R}} K(x) dx = 1 \quad \text{and} \quad \int_{\mathbb{R}} K^2(x) dx = \nu^2.$$

Theorem 6.1. *Assume that (\mathcal{H}_1) to (\mathcal{H}_5) hold and that the sequence $(\varepsilon_{i,j})$ has a finite moment of order > 2 . Then, for any $x \in [-1/2; 1/2]$, we have*

$$(6.5) \quad \lim_{n \rightarrow \infty} \widehat{f}_n(x) = f(x) \quad \text{a.s.}$$

Theorem 6.2. *Assume that (\mathcal{H}_1) to (\mathcal{H}_5) hold and that the sequence $(\varepsilon_{i,j})$ has a finite moment of order > 2 . Then, as soon as the bandwidth (h_n) satisfies $h_n = 1/n^\alpha$ with $\alpha > 1/3$, we have for any $x \in [-1/2; 1/2]$ with $x \neq 0$, the pointwise asymptotic normality*

$$(6.6) \quad \sqrt{nh_n}(\widehat{f}_n(x) - f(x)) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\nu^2}{1+\alpha} \sum_{j=1}^p \frac{\sigma_j^2 \omega_j^2(x)}{a_j^2 (g(\theta_j + x) + g(\theta_j - x))}\right).$$

In addition, for $x = 0$,

$$(6.7) \quad \sqrt{nh_n}(\widehat{f}_n(0) - f(0)) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\nu^2}{1+\alpha} \sum_{j=1}^p \frac{\sigma_j^2 \omega_j^2(0)}{a_j^2 g(\theta_j)}\right).$$

Remark 6.1. *The choice of the weights $\omega_j(x)$ could be important. Intuitively, the asymptotic variances in (6.6) and (6.7) are minimal if for all $1 \leq j \leq p$, $\omega_j(x)$ is inversely proportional to the variance σ_j^2 of the noise. More precisely, the Lagrange Multiplier Theorem gives us the values of the weights for which the asymptotic variances in (6.6) and (6.7) are minimal under the constraint (6.2). They are given, for all $1 \leq j \leq p$ and for all $x \in [-1/2; 1/2]$, by*

$$\omega_j(x) = \frac{m_j(x)}{\sum_{k=1}^p m_k(x)}$$

where

$$m_j(x) = \frac{a_j^2 (g(\theta_j + x) + g(\theta_j - x))}{\sigma_j^2}.$$

For these values, the asymptotic variances in (6.6) and (6.7) are respectively given, for $x \neq 0$, by

$$\frac{\nu^2}{1+\alpha} \left(\sum_{j=1}^p m_j(x) \right)^{-1},$$

and for $x = 0$, by

$$\frac{\nu^2}{1+\alpha} \left(\sum_{j=1}^p \frac{m_j(0)}{2} \right)^{-1}.$$

7. SIMULATIONS

In this section, we illustrate the asymptotic behavior of our estimates on simulated data as well as on real data.

7.1. Simulated data. We consider data simulated according to the model (1.1)

$$Y_{i,j} = a_j f(X_i - \theta_j) + v_j + \varepsilon_{i,j}$$

where $1 \leq j \leq p$ and $1 \leq i \leq n$ with $p = 5$ and $n = 2000$. Moreover, $v = (0, 1/3, -1, 2, -9/10)^T$, $\theta = (0, 1/5, -1/20, -1/7, 1/6)^T$ and $a = (1, -4, 3, -5/2, -2)^T$ and the noise $(\varepsilon_{i,j})$ is a sequence of i.i.d. random variables with $\mathcal{N}(0, 1)$ distribution. The random variables (X_i) are simulated according to the uniform distribution on $[-1/2; 1/2]$ and the regression function, whose $f_1 = 1/2$, is given for all $x \in [-1/2; 1/2]$ by

$$f(x) = \sum_{k=1}^5 \cos(2k\pi x).$$

The simulated data are given in Figure 1. The results of the estimates of the vectors v , θ and a by \hat{v}_n , $\hat{\theta}_n$ and \hat{a}_n are given in Figure 2. The true values are drawn in abscissa and their estimates in ordinate. One can observe that our parametric estimation procedure performs pretty well. For the estimation of the regression function f , we have chosen $\alpha = 9/10$ for the bandwidth sequence (h_n) . Moreover, the kernel K considered is the uniform kernel on $[-1; 1]$ and for all $1 \leq j \leq p$, $\omega_j(x) = 1/p$. The estimation of the regression function f by \hat{f}_n is given on the left side of Figure 3, and the estimation of f by $\hat{f}_{n,1}$ is given on the right side.

One can observe that the estimation of f by \hat{f}_n is better than the one by $\hat{f}_{n,1}$. This is due to the asymptotic variances given in (6.6) and (6.7) of Theorem 6.2. These asymptotic variances are smaller than those of Theorem 3.2 in [1]. In particular, this justifies the choice of taking a weighed version of the Nadaraya-Watson estimator.

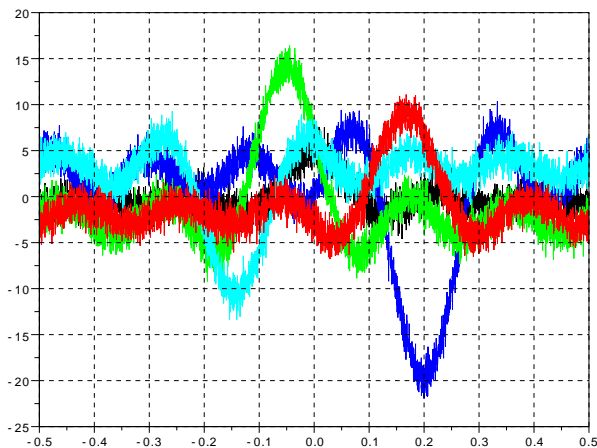
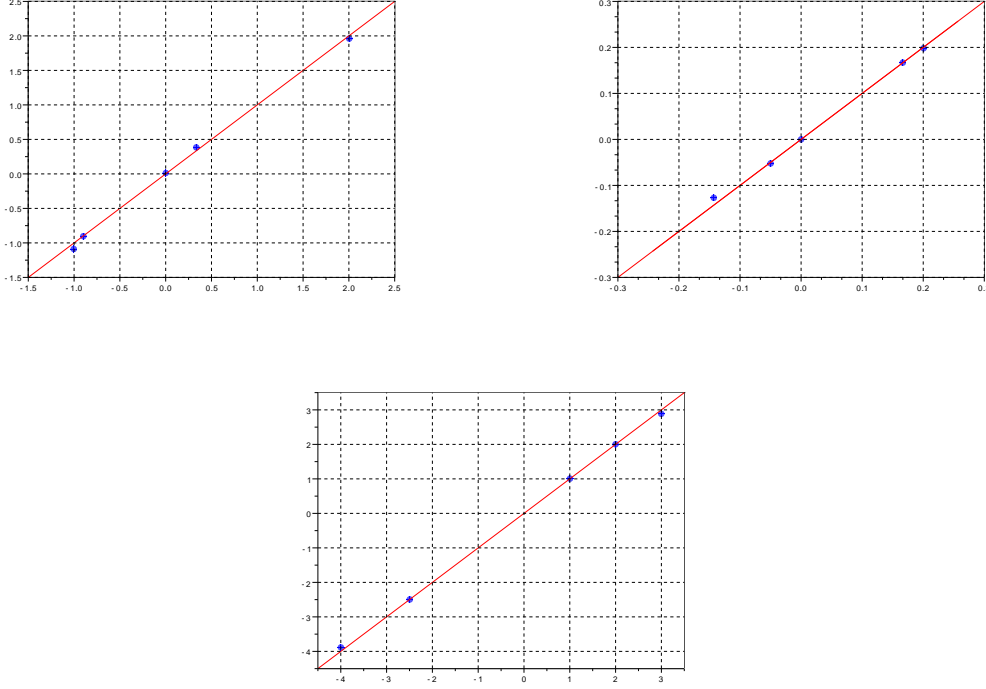
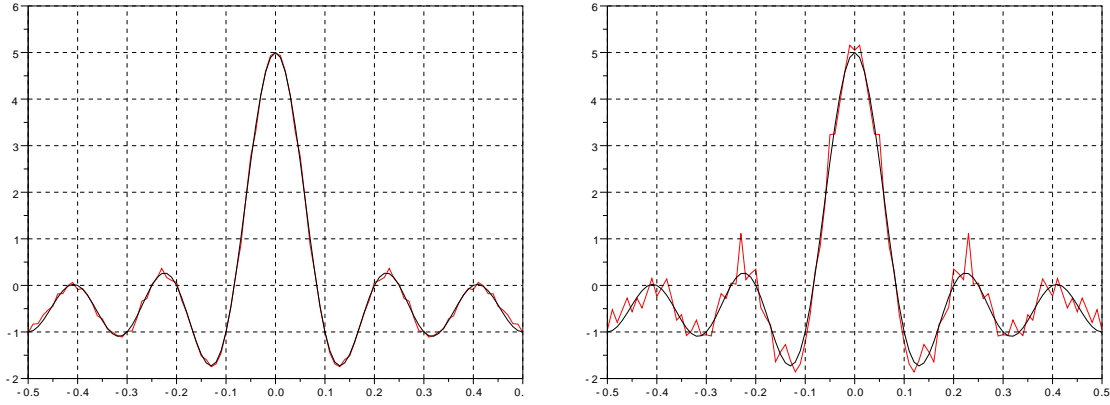


FIGURE 1. Simulated data

FIGURE 2. Estimation of v , θ and a FIGURE 3. Estimation of f by \hat{f}_n and $\hat{f}_{n,1}$

7.2. Modeling of ECG data. This section is devoted to real ECG data. An ECG is a recording of the electrical activity of the heart over a period of time, as detected by electrodes attached to the outer surface of the skin. A typical ECG consists of a P wave, followed by a QRS complex, and a T wave. Here, we consider two sets of ECG data, one corresponding to a healthy heart and the other to a heart having arrhythmia. These data are extracted from the MIT-BIH Database, and they are represented respectively in Figure 4 and in Figure 5.

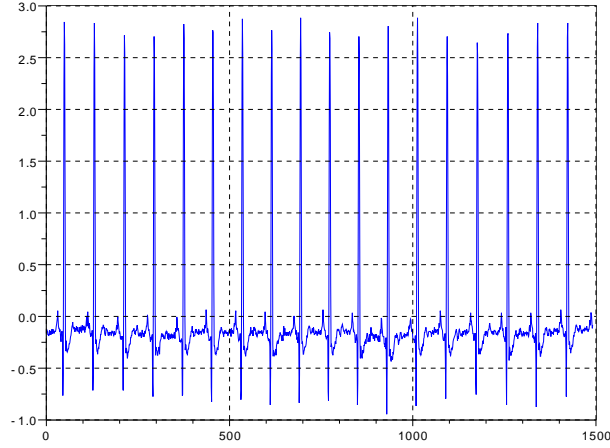


FIGURE 4. ECG of a healthy patient

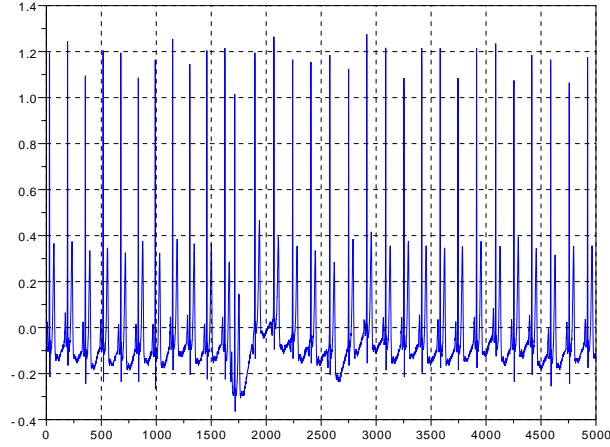


FIGURE 5. ECG of a patient having arrhythmia

For each ECG recording, one considers that the heart cycle of interest, that is to say the cycle PQRST, is roughly the same at each beat. One also considers that every heart cycle is noisy and that the white noise is independent of the typical shape we want to estimate. After an appropriate segmentation of the ECGs, one observes signals of the same length such that each of them contains a unique PQRST cycle. The segmentation of the ECGs is done by detecting the maximum of each QRS complex and centering the segments around this maximum. It is very important to have segments of the same length in order to ensure periodicity. A well-adapted method for the segmentation is the one proposed by Gasser and Kneip [8]. After the segmentation of the two ECGs, we obtain $p = 18$ and every segment of length $n = 83$ for the healthy heart and $p = 15$ and $n = 91$ for the ill heart. Then, our goal

is to estimate the typical shape of each ECG, corresponding to the function f in the model (1.1). Firstly, the estimation for the healthy heart is going to show that our model is well adapted for the problem of modelling an ECG signal. Indeed, for the healthy heart, a good approximation of the heart cycle is to take the average of the p different signals, whereas the different parameters v , θ and a of the model (1.1) are trivial. The estimation of the typical shape of the ECG for the healthy heart by our procedure is on the right hand-side of Figure 6, whereas the left-side shows one on the original ECG signal. The comparison between the two plots shows that our estimation procedure performs pretty well and that our estimation procedure can be useful for the modeling of an ECG signal.

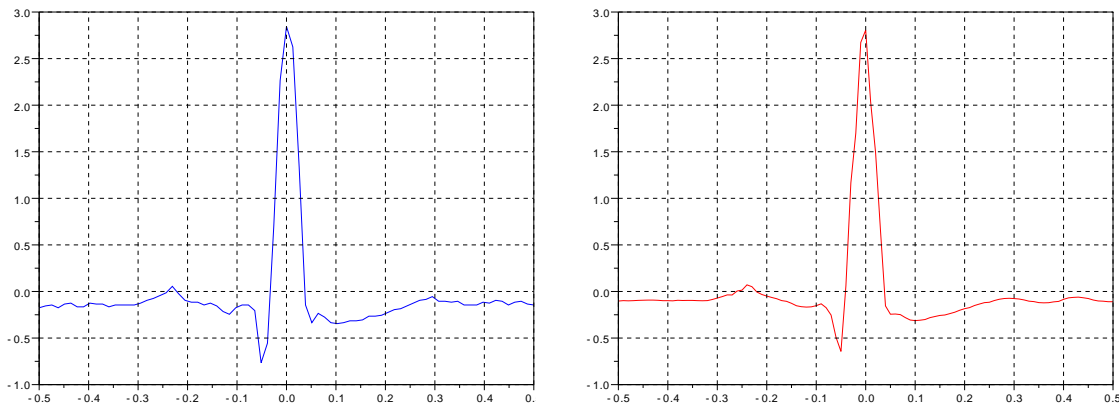


FIGURE 6. One signal and its reconstruction

Secondly, we want to estimate the typical shape of the nonhealthy heart, plotted in Figure 5. One see that the electric activity is more irregular than for the healthy heart, and a simple averaging may lead to a mean cycle that does not correspond to the typical shape of the ECG. More precisely, we suppose that the model (1.1)

$$Y_{i,j} = a_j f(X_i - \theta_j) + v_j + \varepsilon_{i,j},$$

fits the data. The segmentation allows us to have a common shape function f 1-periodic. Moreover, f is nonsymmetric, but we already saw in Remark 4.3 that our procedure still works for nonsymmetric shape function. The parameters a , θ and v correspond to the deformation due to the arrhythmia relative to the common shape f we want to estimate. For more accuracy, we have to choose one of the p curves as a reference, that is to say where for one $1 \leq j^* \leq p$, $a_{j^*} = 1$, $\theta_{j^*} = 0$ and $v_{j^*} = 0$. For this choice, we consider a criterion of residual variance. More precisely, we first consider the model (1.1) with $a_1 = 1$, $\theta_1 = 0$ and $v_1 = 0$. From this model, we apply our procedure to estimate the different parameters a , θ and v and the shape function f respectively by \hat{a}_n , $\hat{\theta}_n$, \hat{v}_n and \hat{f}_n . With these estimates, we then calculate the

vector $\hat{\sigma}_{n,1}^2$ whose j -th component $(\hat{\sigma}_{n,1}^2)_j$ is defined by

$$(\hat{\sigma}_{n,1}^2)_j = \frac{1}{n} \sum_{i=1}^n \left(Y_{i,j} - \hat{a}_{n,j} \hat{f}_n(X_i - \hat{\theta}_{n,j}) - \hat{v}_{n,j} \right)^2.$$

Then, we make use of the same procedure by changing the curve of reference, and we finally obtain p vectors of length p

$$\hat{\sigma}_{n,1}^2, \dots, \hat{\sigma}_{n,p}^2.$$

Finally, the choice of the curve of reference for the modeling of the ECG signal is given by taking

$$j^* = \arg \min_{1 \leq j \leq p} \|\hat{\sigma}_{n,j}^2\|_1$$

where $\|\cdot\|_1$ corresponds to the l^1 -norm. Therefore, we model the ECG by

$$Y_{i,j} = a_j f(X_i - \theta_j) + v_j + \varepsilon_{i,j}$$

where

$$1 \leq i \leq n, \quad 1 \leq j \leq p \quad \text{and} \quad a_{j^*} = 1, \theta_{j^*} = 0, v_{j^*} = 0.$$

On our data set, the implementation of this method shows that $j^* = 3$ and

$$\|\hat{\sigma}_{n,j^*}^2\|_1 = 0.6095.$$

The result for the estimation of the typical shape f is given in Figure 7. On the right side of Figure 7 one can compare the estimation of f obtained with our estimate \hat{f}_n and with the simple average signal. One can observe that our estimation procedure is better because the P wave is well estimated.

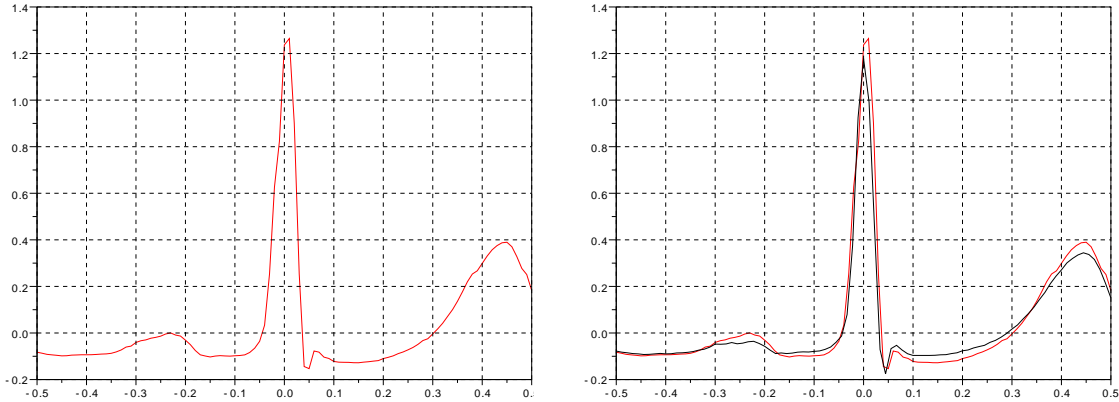


FIGURE 7. Estimation of f by \hat{f}_n (in red) and with the mean average signal (in black)

8. PROOFS OF THE PARAMETRIC RESULTS

8.1. Proof of Theorem 3.1. The convergences (3.3) and (3.4) follow from the standard law of large numbers and the standard central limit theorem for martingales with independent increments. Moreover, we immediately deduce (3.5) from Theorem 2.1 of [3]. \square

8.2. Proof of Theorem 4.1. The result follows from Theorem 2.1 of [1]. \square

8.3. Proof of Theorem 4.2. Our goal is to apply Theorem 2.1 page 330 of Kushner and Yin [12]. First of all, as $\gamma_n = 1/n$, the condition on the decreasing step is satisfied. Moreover, we already saw that $\widehat{\theta}_n$ converges almost surely to θ . Consequently, all the local assumptions of Theorem 2.1 of Kushner and Yin [12] are satisfied. In addition, it is not hard to see from (4.7) that

$$\mathbb{E}[T_{n+1}|\mathcal{F}_n] = \phi(\widehat{\theta}_n) \quad \text{a.s.}$$

Moreover, the function ϕ is continuously differentiable. Hence, $\phi(\theta) = 0$ and $D\phi(\theta)$ is the square diagonal matrix defined by

$$D\phi(\theta) = -2\pi f_1 \text{diag}(a_1, \dots, a_p).$$

By noting I_p the identity matrix, the condition $4\pi|f_1|\min_{1 \leq j \leq p}|a_j| > 1$ implies that

$$D\phi(\theta) + \frac{1}{2}I_p$$

is a negative-definite matrix. Furthermore, we have that, for all $1 \leq k, l \leq p$,

$$\mathbb{E}[\text{sign}(a_k f_1)T_{n+1,k} \text{sign}(a_l f_1)T_{n+1,l}|\mathcal{F}_n] = \varphi(\widehat{\theta}_n)_{k,l} \quad \text{a.s.}$$

which leads to

$$\lim_{n \rightarrow \infty} \mathbb{E}[\text{sign}(a_k f_1)T_{n+1,k} \text{sign}(a_l f_1)T_{n+1,l}|\mathcal{F}_n] = \varphi(\theta)_{k,l} \quad \text{a.s.}$$

Consequently, if we are able to prove that the sequence (W_n) given by

$$W_n = \frac{\|\widehat{\theta}_n - \theta\|^2}{\gamma_n}$$

is tight, then we shall deduce from Theorem 2.1 of [12] that

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}_p(0, \Sigma(\theta))$$

where for all $1 \leq k, l \leq p$,

$$\Sigma(\theta)_{k,l} = \varphi(\theta)_{k,l} \int_0^{+\infty} \exp\left((1 - 2\pi|f_1|(|a_k| + |a_l|))t\right) dt = \frac{\varphi(\theta)_{k,l}}{2\pi|f_1|(|a_k| + |a_l|) - 1}.$$

Therefore, it remains to prove the tightness of the sequence (W_n) . Let (V_n) be the sequence defined, for all $n \geq 1$, by

$$(8.1) \quad V_n = \|\widehat{\theta}_n - \theta\|^2,$$

and T'_n the sequence defined, for all $1 \leq j \leq p$, by

$$(8.2) \quad T'_{n,j} = \text{sign}(a_j f_1) T_{n,j}.$$

Then, we clearly have

$$\begin{aligned} V_{n+1} &= \|\widehat{\theta}_{n+1} - \theta\|^2 \\ &= \|\pi_{K^p}(\widehat{\theta}_n + \gamma_{n+1} T'_{n+1}) - \theta\|^2 \\ &= \|\pi_{K^p}(\widehat{\theta}_n + \gamma_{n+1} T'_{n+1}) - \pi_{K^p}(\theta)\|^2 \\ &\leq \|\widehat{\theta}_n + \gamma_{n+1} T'_{n+1} - \theta\|^2 \end{aligned}$$

as $\pi_{K^p} = (\pi_K, \dots, \pi_K)^T$ is a Lipschitz function. It follows that

$$V_{n+1} \leq V_n + \gamma_{n+1}^2 \|T'_{n+1}\|^2 + 2\gamma_{n+1} \langle \widehat{\theta}_n - \theta, T'_{n+1} \rangle \quad \text{a.s.}$$

By taking expectation in the previous inequality, we obtain that there exists a constant $M > 0$ such that

$$(8.3) \quad \mathbb{E}[V_{n+1} | \mathcal{F}_n] \leq V_n + \gamma_{n+1}^2 M + 2\gamma_{n+1} \langle \widehat{\theta}_n - \theta, \mathbb{E}[T'_{n+1} | \mathcal{F}_n] \rangle \quad \text{a.s.}$$

Moreover, (4.7) together with (8.2) lead to

$$(8.4) \quad \mathbb{E}[T'_{n+1} | \mathcal{F}_n] = S_p(a) \phi(\widehat{\theta}_n),$$

where

$$S_p(a) = \text{diag}(\text{sign}(a_1 f_1), \dots, \text{sign}(a_p f_1)).$$

Hence, we deduce from (8.3) and (8.4) that

$$(8.5) \quad \mathbb{E}[W_{n+1} | \mathcal{F}_n] \leq \frac{V_n}{\gamma_{n+1}} + \gamma_{n+1} M + 2 \langle \widehat{\theta}_n - \theta, S_p(a) \phi(\widehat{\theta}_n) \rangle \quad \text{a.s.}$$

Moreover, a Taylor expansion of ϕ allows us to write

$$\begin{aligned} \langle \widehat{\theta}_n - \theta, S_p(a) \phi(\widehat{\theta}_n) \rangle &= \langle \widehat{\theta}_n - \theta, 2\pi f_1 S_p(a) \text{diag}(a_1, \dots, a_p) (\theta - \widehat{\theta}_n) \rangle \\ (8.6) \quad &+ f_1 \langle \widehat{\theta}_n - \theta, S_p(a) \text{diag}(a_1, \dots, a_p) \mathcal{V}(\widehat{\theta}_n) (\theta - \widehat{\theta}_n) \rangle, \end{aligned}$$

where for all $t \neq \theta$,

$$\mathcal{V}(t) = \text{diag}\left(\frac{\sin(2\pi(\theta_1 - t_1)) - 2\pi(\theta_1 - t_1)}{\theta_1 - t_1}, \dots, \frac{\sin(2\pi(\theta_p - t_p)) - 2\pi(\theta_p - t_p)}{\theta_p - t_p}\right).$$

Moreover, the equality

$$f_1 S_p(a) \text{diag}(a_1, \dots, a_p) = L(a),$$

where

$$L(a) = \text{diag}(|f_1 a_1|, \dots, |f_1 a_p|),$$

together with (8.6) lead to

$$\begin{aligned} \langle \hat{\theta}_n - \theta, S_p(a)\phi(\hat{\theta}_n) \rangle &= -2\pi (\hat{\theta}_n - \theta)^T L(a) (\hat{\theta}_n - \theta) \\ &\quad - (\hat{\theta}_n - \theta)^T L(a) \mathcal{V}(\hat{\theta}_n) (\hat{\theta}_n - \theta). \end{aligned}$$

Hence, (8.5) can be rewritten as

$$\begin{aligned} (8.7) \quad \mathbb{E}[W_{n+1}|\mathcal{F}_n] &\leq (1 + \gamma_n) W_n + \gamma_n M - 4\pi (\hat{\theta}_n - \theta)^T L(a) (\hat{\theta}_n - \theta) \\ &\quad - 2 (\hat{\theta}_n - \theta)^T L(a) \mathcal{V}(\hat{\theta}_n) (\hat{\theta}_n - \theta). \end{aligned}$$

Moreover, as

$$L(a) \geq \min_{1 \leq j \leq p} |a_j f_1| I_p,$$

we deduce from (8.7) that

$$(8.8) \quad \mathbb{E}[W_{n+1}|\mathcal{F}_n] \leq W_n + 2q\gamma_n W_n + M\gamma_n - 2 (\hat{\theta}_n - \theta)^T L(a) (\hat{\theta}_n - \theta),$$

where

$$2q = 1 - 4\pi |f_1| \min_{1 \leq j \leq p} |a_j|,$$

which means that $q < 0$. By the continuity of the function \mathcal{V} , one can find $0 < \varepsilon < 1/2$ such that, if $\|\hat{\theta}_n - \theta\| < \varepsilon$,

$$(8.9) \quad \frac{q}{2|f_1| \min_{1 \leq j \leq p} |a_j|} I_p < \mathcal{V}(\hat{\theta}_n) < 0.$$

Moreover, let A_n and B_n be the sets $A_n = \{\|\hat{\theta}_n - \theta\| \leq \varepsilon\}$ and

$$B_n = \bigcap_{k=m}^n A_k$$

with $1 \leq m \leq n$. Then, it follows from (8.9) that

$$(8.10) \quad 0 < -2|f_1| \min_{1 \leq j \leq p} |a_j| \mathcal{V}(\hat{\theta}_n) I_{B_n} < -\left(\frac{q}{2}\right) I_p I_{B_n}.$$

where I_p is the identity matrix. Hence, it follows from the conjunction of (8.8) and (8.10) that for all $n \geq m$,

$$\begin{aligned} (8.11) \quad \mathbb{E}[W_{n+1} I_{B_n} | \mathcal{F}_n] &\leq W_n I_{B_n} + 2\gamma_n W_n q I_{B_n} - q\gamma_n W_n I_{B_n} + \gamma_n M, \\ &\leq W_n I_{B_n} (1 + q\gamma_n) + \gamma_n M. \end{aligned}$$

Since $B_{n+1} = B_n \cap A_{n+1}$, $B_{n+1} \subset B_n$, and we obtain by taking the expectation on both sides of (8.11) that for all $n \geq m$,

$$(8.12) \quad \mathbb{E}[W_{n+1} I_{B_{n+1}}] \leq (1 + q\gamma_n) \mathbb{E}[W_n I_{B_n}] + \gamma_n M.$$

Finally, following the same lines as in the proof of Theorem 2.2 in [1], we obtain that for all $\xi > 0$, it exists $K > 0$ such that for m large enough,

$$\sup_{n \geq m} \mathbb{P}(W_n > K) < \xi$$

which implies the tightness of (W_n) and completes the proof of Theorem 4.2. \square

8.4. Proof of Theorem 4.3. The vectorial law of iterated logarithm given by (4.10) follows from Theorem 1 of [15], while the vectorial quadratic strong law given by (4.12) can be obtained from Theorem 3 of [15]. \square

8.5. Proof of Theorem 5.1. Recall that for all $1 \leq j \leq p$,

$$\hat{a}_{n,j} = \frac{1}{nf_1} \sum_{i=1}^n \frac{\cos(2\pi(X_i - \hat{\theta}_{i-1,j}))}{g(X_i)} Y_{i,j}.$$

Then, it is clear that

$$\hat{a}_n = \frac{1}{nf_1} \sum_{i=1}^n C(X_i, \hat{\theta}_{i-1}) Y_i,$$

where

$$Y_i = \begin{pmatrix} Y_{i,1} \\ \vdots \\ Y_{i,p} \end{pmatrix}.$$

We also have the decomposition

$$(8.13) \quad \hat{a}_n - a = \frac{1}{nf_1} S_n(a) + \frac{1}{nf_1} R_n(a),$$

with

$$S_n(a) = \sum_{i=1}^n (C(X_i, \theta) Y_i - f_1 a),$$

and the remainder

$$R_n(a) = \sum_{i=1}^n \left(C(X_i, \hat{\theta}_{i-1}) - C(X_i, \theta) \right) Y_i.$$

Moreover, for all $1 \leq j \leq p$,

$$(8.14) \quad R_{n,j}(a) = a_j R_{n,j}^1(a) + v_j R_{n,j}^2(a) + R_{n,j}^3(a),$$

where

$$R_{n,j}^1(a) = \sum_{i=1}^n \frac{\Delta c_{i,j}}{g(X_i)} f(X_i - \theta_j), \quad R_{n,j}^2(a) = \sum_{i=1}^n \frac{\Delta c_{i,j}}{g(X_i)}, \quad R_{n,j}^3(a) = \sum_{i=1}^n \frac{\Delta c_{i,j}}{g(X_i)} \varepsilon_{i,j},$$

and

$$\Delta c_{i,j} = \cos\left(2\pi(X_i - \hat{\theta}_{i-1,j})\right) - \cos\left(2\pi(X_i - \theta_j)\right).$$

Firstly, since

$$\begin{aligned}\mathbb{E}\left[\frac{\Delta c_{i,j}}{g(X_i)}|\mathcal{F}_{i-1}\right] &= \int_{-1/2}^{1/2} \cos\left(2\pi(x - \widehat{\theta}_{i-1,j})\right) dx - \int_{-1/2}^{1/2} \cos(2\pi(x - \theta_j)) dx \\ &= 0,\end{aligned}$$

the sequence $(R_{n,j}^2(a))$ is a square integrable martingale whose predictable variation is given by

$$\langle R_j^2(a) \rangle_n = \sum_{i=1}^n \mathbb{E}\left[\frac{\Delta c_{i,j}^2}{g^2(X_i)}|\mathcal{F}_{i-1}\right].$$

Then, as \cos is a Lipschitz function, we have

$$|\Delta c_{i,j}| \leq |\widehat{\theta}_{i-1,j} - \theta_j|.$$

Consequently, since g does not vanish on $[-1/2; 1/2]$, there exists a constant $C > 0$ such that

$$(8.15) \quad \mathbb{E}\left[\frac{\Delta c_{i,j}^2}{g^2(X_i)}|\mathcal{F}_{i-1}\right] \leq C \left(\widehat{\theta}_{i-1,j} - \theta_j\right)^2.$$

Then, it follows from (4.12) together with the previous inequality (8.15) that

$$(8.16) \quad \langle R_j^2(a) \rangle_n = \mathcal{O}(\log(n)) \quad \text{a.s.}$$

Therefore, we deduce from the strong law of large numbers for martingales given e.g. by Theorem 1.3.15 of [6] that, for all $1 \leq j \leq p$,

$$(8.17) \quad R_{n,j}^2(a) = o(\log(n)) \quad \text{a.s.}$$

Moreover, $(R_{n,j}^3(a))$ is also a square integrable martingale whose predictable variation is given by

$$\langle R_j^3(a) \rangle_n = \sigma_j^2 \langle R_j^2(a) \rangle_n.$$

Then, we immediately deduce from (8.16) that

$$(8.18) \quad \langle R_j^3(a) \rangle_n = \mathcal{O}(\log(n)) \quad \text{a.s.},$$

and from the strong law of large numbers for martingales that, for all $1 \leq j \leq p$,

$$(8.19) \quad R_{n,j}^3(a) = o(\log(n)) \quad \text{a.s.}$$

Afterwards, for all $1 \leq j \leq p$, with the change of variables $u = x - \theta_j$,

$$\begin{aligned}\mathbb{E}\left[\frac{\Delta c_{i,j}}{g(X_i)}f(X_i - \theta_j)|\mathcal{F}_{i-1}\right] &= \int_{-1/2}^{1/2} \left(\cos(2\pi(x - \widehat{\theta}_{i-1,j})) - \cos(2\pi(x - \theta_j))\right) f(x - \theta_j) dx \\ &= \int_{-1/2-\theta_j}^{1/2-\theta_j} \left(\cos(2\pi(u + \theta_j - \widehat{\theta}_{i-1,j})) - \cos(2\pi u)\right) f(u) du.\end{aligned}$$

Then, the elementary trigonometric equality

$$\cos(2\pi(u + \theta_j - \widehat{\theta}_{i-1,j})) = \cos(2\pi u) \cos(2\pi(\theta_j - \widehat{\theta}_{i-1,j})) - \sin(2\pi u) \sin(2\pi(\theta_j - \widehat{\theta}_{i-1,j})),$$

the symmetry and the periodicity of the function f lead to

$$\mathbb{E}\left[\frac{\Delta c_{i,j}}{g(X_i)}f(X_i - \theta_j)|\mathcal{F}_{i-1}\right] = f_1 \left(\cos(2\pi(\theta_j - \widehat{\theta}_{i-1,j})) - 1\right) \quad \text{a.s.}$$

Moreover, for all $|x| < 1/2$, we have

$$|\cos(2\pi x) - 1| \leq 2\pi^2 x^2,$$

which implies that

$$(8.20) \quad |\mathbb{E}\left[\frac{\Delta c_{i,j}}{g(X_i)} f(X_i - \theta_j) | \mathcal{F}_{i-1}\right]| \leq 2\pi^2 \left(\theta_j - \widehat{\theta}_{i-1,j}\right)^2 \quad \text{a.s.}$$

Moreover, we have the decomposition

$$R_{n,j}^1(a) = A_{n,j}(a) + B_{n,j}(a),$$

where

$$A_{n,j}(a) = \sum_{i=1}^n \left(\frac{\Delta c_{i,j}}{g(X_i)} f(X_i - \theta_j) - \mathbb{E}\left[\frac{\Delta c_{i,j}}{g(X_i)} f(X_i - \theta_j) | \mathcal{F}_{i-1}\right] \right),$$

and

$$B_{n,j}(a) = \sum_{i=1}^n \mathbb{E}\left[\frac{\Delta c_{i,j}}{g(X_i)} f(X_i - \theta_j) | \mathcal{F}_{i-1}\right].$$

It follows one again from the quadratic strong law (4.12) together with (8.20) that, for all $1 \leq j \leq p$,

$$(8.21) \quad B_{n,j}(a) = \mathcal{O}(\log(n)) \quad \text{a.s.}$$

Moreover, for all $1 \leq j \leq p$, $(A_{n,j}(a))$ is a square integrable martingale whose predictable variation $\langle A_j(a) \rangle_n$ satisfy

$$\langle A_j(a) \rangle_n \leq \sum_{i=1}^n \mathbb{E}\left[\frac{\Delta c_{i,j}^2}{g(X_i)^2} f^2(X_i - \theta_j) | \mathcal{F}_{i-1}\right].$$

As the shape function f is bounded, we deduce from (4.12) together with (8.15) that

$$\langle A_j(a) \rangle_n = \mathcal{O}(\log(n)) \quad \text{a.s.}$$

Therefore, we can conclude from the strong law of large numbers for martingales that, for all $1 \leq j \leq p$,

$$(8.22) \quad A_{n,j}(a) = o(\log(n)) \quad \text{a.s.}$$

Finally, we infer from (8.17), (8.19) together with (8.21) and (8.22) that, for all $1 \leq j \leq p$,

$$(8.23) \quad R_{n,j}(a) = \mathcal{O}(\log(n)) \quad \text{a.s.}$$

Hence, one obtain from (8.13) that

$$(8.24) \quad \widehat{a}_n - a = \frac{1}{nf_1} S_n(a) + \mathcal{O}\left(\frac{\log(n)}{n}\right) \quad \text{a.s.}$$

Secondly, since

$$\mathbb{E}[C(X_i, \theta) Y_i | \mathcal{F}_{i-1}] = f_1 a,$$

the sequence $(S_n(a))$ is a vectorial martingale with independent increments. For all $n \geq 1$, its predictable variation $\langle S(a) \rangle_n$ is given by

$$\begin{aligned} \langle S(a) \rangle_n &= \sum_{i=1}^n \mathbb{E}[(C(X_i, \theta)Y_i - f_1 a)(C(X_i, \theta)Y_i - f_1 a)^T | \mathcal{F}_{i-1}] \\ &= \sum_{i=1}^n \text{Cov}(C(X_i, \theta)Y_i | \mathcal{F}_{i-1}) \end{aligned}$$

Then, it is clear that

$$\lim_{n \rightarrow +\infty} \frac{\langle S(a) \rangle_n}{n} = \Gamma(a) \quad \text{a.s.}$$

where $\Gamma(a)$ is given by (5.7). Consequently, (5.5) follows from the law of large numbers for martingales and (5.6) follows from the central limit theorem for martingales and Slutsky's lemma, while one can obtain (5.8) from Theorem 2.1 of [3]. \square

9. PROOFS OF THE NONPARAMETRIC RESULTS

9.1. Proof of Theorem 6.1. For $x \in [-1/2; 1/2]$, denote by $\tilde{f}_{n,j}(x)$ the sequence defined for $n \geq 1$ and $1 \leq j \leq p$, by

$$(9.1) \quad \tilde{f}_{n,j}(x) = \hat{a}_{n,j} \hat{f}_{n,j}(x).$$

We can rewrite

$$(9.2) \quad \tilde{f}_{n,j}(x) = \tilde{f}_{n,j}^1(x) + \tilde{f}_{n,j}^2(x),$$

where

$$\tilde{f}_{n,j}^1(x) = \frac{\sum_{i=1}^n (W_{i,j}(x) + W_{i,j}(-x)) (Y_{i,j} - v_j)}{\sum_{i=1}^n (W_{i,j}(x) + W_{i,j}(-x))},$$

and

$$\tilde{f}_{n,j}^2(x) = \frac{\sum_{i=1}^n (W_{i,j}(x) + W_{i,j}(-x)) (v_j - \hat{v}_{i-1,j})}{\sum_{i=1}^n (W_{i,j}(x) + W_{i,j}(-x))}.$$

On the one hand, it follows from Theorem 3.1 of [1] that for any $x \in [-1/2; 1/2]$,

$$\lim_{n \rightarrow +\infty} \tilde{f}_{n,j}^1(x) = a_j f(x) \quad \text{a.s.}$$

On the other hand, the almost sure convergence of $\hat{v}_{i-1,j}$ to v_j as i goes to infinity implies by Toeplitz lemma, that for any $x \in [-1/2; 1/2]$,

$$\lim_{n \rightarrow +\infty} \tilde{f}_{n,j}^2(x) = 0 \quad \text{a.s.}$$

Hence, one can conclude that

$$(9.3) \quad \lim_{n \rightarrow \infty} \tilde{f}_{n,j}(x) = a_j f(x) \quad \text{a.s.}$$

Consequently, as $\hat{a}_{n,j}$ converges almost surely to $a_j \neq 0$ as n goes to infinity, it follows that

$$(9.4) \quad \lim_{n \rightarrow \infty} \hat{f}_{n,j}(x) = f(x) \quad \text{a.s.}$$

Finally, (6.1) with (9.4) allow us to conclude the proof of Theorem 6.1. \square

9.2. Proof of Theorem 6.2. We shall now proceed to the proof of the asymptotic normality of \widehat{f}_n . We have, for all $x \in [-1/2; 1/2]$,

$$\begin{aligned}
 \widehat{f}_n(x) - f(x) &= \sum_{j=1}^p \omega_j(x) \left(\widehat{f}_{n,j}(x) - f(x) \right) \\
 (9.5) \quad &= \sum_{j=1}^p \omega_j(x) \frac{\mathcal{M}_{n,j}(x) + \mathcal{P}_{n,j}(x) + \mathcal{Q}_{n,j}(x) + \mathcal{R}_{n,j}(x) + \mathcal{S}_{n,j}(x)}{n\mathcal{G}_{n,j}(x)}
 \end{aligned}$$

where

$$\begin{aligned}
 \mathcal{G}_{n,j}(x) &= \widehat{a}_{n,j} (\widehat{g}_{n,j}(x) + \widehat{g}_{n,j}(-x)), \\
 \mathcal{M}_{n,j}(x) &= M_{n,j}(x) + M_{n,j}(-x), \\
 \mathcal{P}_{n,j}(x) &= P_{n,j}(x) + P_{n,j}(-x), \\
 \mathcal{Q}_{n,j}(x) &= Q_{n,j}(x) + Q_{n,j}(-x), \\
 \mathcal{R}_{n,j}(x) &= R_{n,j}(x) + R_{n,j}(-x), \\
 \mathcal{S}_{n,j}(x) &= S_{n,j}(x) + S_{n,j}(-x),
 \end{aligned}$$

with $\widehat{g}_{n,j}(x)$, $M_{n,j}(x)$, $P_{n,j}(x)$, $Q_{n,j}(x)$, $R_{n,j}(x)$ and $S_{n,j}(x)$ given by

$$\begin{aligned}
 \widehat{g}_{n,j}(x) &= \frac{1}{n} \sum_{i=1}^n W_{i,j}(x), \\
 M_{n,j}(x) &= \sum_{i=1}^n W_{i,j}(x) \varepsilon_{i,j}, \\
 P_{n,j}(x) &= \widehat{a}_{n,j} \sum_{i=1}^n W_{i,j}(x) \left(f(X_i - \widehat{\theta}_{i-1,j}) - f(x) \right), \\
 Q_{n,j}(x) &= \widehat{a}_{n,j} \sum_{i=1}^n W_{i,j}(x) \left(f(X_i - \theta_j) - f(X_i - \widehat{\theta}_{i-1,j}) \right), \\
 R_{n,j}(x) &= (a_j - \widehat{a}_{n,j}) \sum_{i=1}^n W_{i,j}(x) f(X_i - \theta_j), \\
 S_{n,j}(x) &= \sum_{i=1}^n W_{i,j}(x) (v_j - \widehat{v}_{i-1,j}).
 \end{aligned}$$

Firstly, (6.28) of [1] together with the almost sure convergence of \widehat{a}_n to a as n goes to infinity, lead to

$$(9.6) \quad \lim_{n \rightarrow +\infty} \mathcal{G}_{n,j}(x) = a_j (g(\theta_j + x) + g(\theta_j - x)) \quad \text{a.s.}$$

In addition, we obtain from (6.32) and (6.35) of [1] that, for $\alpha > 1/3$,

$$(9.7) \quad \mathcal{P}_{n,j}^2(x) = o(n^{1+\alpha}) \quad \text{a.s.},$$

$$(9.8) \quad \mathcal{Q}_{n,j}^2(x) = o(n^{1+\alpha}) \quad \text{a.s.}$$

Hence, for all $x \in [-1/2, 1/2]$, we find that

$$(9.9) \quad \lim_{n \rightarrow \infty} \sqrt{\frac{h_n}{n}} \sum_{j=1}^p \omega_j(x) \frac{\mathcal{P}_{n,j}(x) + \mathcal{Q}_{n,j}(x)}{\mathcal{G}_{n,j}(x)} = 0 \quad \text{a.s.}$$

Secondly, as the shape function f is bounded, it follows that

$$R_{n,j}(x) = \mathcal{O} \left(|a_j - \hat{a}_{n,j}| \sum_{i=1}^n W_{i,j}(x) \right) \quad \text{a.s.}$$

Hence,

$$\sum_{i=1}^n W_{i,j}(x) = \mathcal{O}(n) \quad \text{a.s.},$$

ensures that

$$(9.10) \quad R_{n,j}(x) = \mathcal{O}(n|a_j - \hat{a}_{n,j}|) \quad \text{a.s.}$$

In addition, we can deduce from (8.24) that, for all $1 \leq j \leq p$,

$$|a_j - \hat{a}_{n,j}| = \mathcal{O} \left(\sqrt{\frac{\log(n)}{n}} \right) \quad \text{a.s.}$$

which via (9.10), leads to

$$(9.11) \quad R_{n,j}(x) = \mathcal{O} \left(\sqrt{n \log(n)} \right) \quad \text{a.s.}$$

Consequently,

$$(9.12) \quad \mathcal{R}_{n,j}^2(x) = o(n^{1+\alpha}) \quad \text{a.s.}$$

Thirdly, we have the following inequality

$$(9.13) \quad |S_{n,j}(x)| \leq \Lambda_{n,j}(x) + \Sigma_{n,j}(x)$$

where

$$\Lambda_{n,j}(x) = \sum_{i=1}^n \mathcal{L}_{i,j}(W_{i,j}(x) - \mathbb{E}[W_{i,j}(x)|\mathcal{F}_{i-1}])$$

and

$$\Sigma_{n,j}(x) = \sum_{i=1}^n \mathcal{L}_{i,j} \mathbb{E}[W_{i,j}(x)|\mathcal{F}_{i-1}]$$

with $\mathcal{L}_{i,j} = |v_j - \widehat{v}_{i-1,j}|$. We deduce from (6.34) of [1] together with the Cauchy-Schwarz inequality and the quadratic strong law given by (3.5) that

$$(9.14) \quad \Sigma_{n,j}(x) = \mathcal{O} \left(\sqrt{n} \left(\sum_{i=1}^n \mathcal{L}_{i,j}^2 \right)^{1/2} \right) = \mathcal{O} \left(\sqrt{n \log(n)} \right) \quad \text{a.s.}$$

Moreover, the sequence $(\Lambda_{n,j}(x))$ is a martingale whose predictable variation is given by

$$\langle \Lambda_j(x) \rangle_n = \mathcal{O} \left(\sum_{i=1}^n \mathcal{L}_{i,j}^2 \mathbb{E}[W_{i,j}^2(x) | \mathcal{F}_{i-1}] \right) \quad \text{a.s.}$$

Consequently, we obtain one again from the quadratic strong law (3.5) that

$$(9.15) \quad \langle \Lambda(x) \rangle_{n,j} = \mathcal{O} \left(n^\alpha \sum_{i=1}^n \mathcal{L}_{i,j}^2 \right) = \mathcal{O} (n^\alpha \log(n)) \quad \text{a.s.}$$

which allows us to show, from the strong law of large numbers for martingales that for any $\gamma > 0$,

$$(9.16) \quad \Lambda_{n,j}^2(x) = o(n^\alpha \log(n)^{2+\gamma}) \quad \text{a.s.}$$

Therefore,

$$(9.17) \quad \begin{aligned} \mathcal{S}_{n,j}^2(x) &= o(n^\alpha \log(n)^{2+\gamma}) + \mathcal{O}(n \log(n)) \\ &= o(n^{1+\alpha}) \quad \text{a.s.} \end{aligned}$$

We are now in position to study the asymptotic behavior of the dominating term $\mathcal{M}_{n,j}(x)$. For all $x \in [-1/2; 1/2]$ and for all $1 \leq j \leq p$, the sequence $(\mathcal{M}_{n,j}(x))$ is a square-integrable martingale whose predictable variation is given by

$$\langle \mathcal{M}_j(x) \rangle_n = \sigma_j^2 \sum_{i=1}^n \mathbb{E} [(W_{i,j}(x) + W_{i,j}(-x))^2 | \mathcal{F}_{i-1}].$$

We deduce from (6.37) of [1] that we have, for $x \neq 0$,

$$(9.18) \quad \lim_{n \rightarrow \infty} \frac{\langle \mathcal{M}_j(x) \rangle_n}{n^{1+\alpha}} = \frac{\sigma_j^2 \nu^2}{1+\alpha} (g(\theta_j + x) + g(\theta_j - x)).$$

and from (6.38) of [1] that, for $x = 0$,

$$(9.19) \quad \lim_{n \rightarrow \infty} \frac{\langle \mathcal{M}_j(0) \rangle_n}{n^{1+\alpha}} = 4 \frac{\sigma_j^2 \nu^2}{1+\alpha} g(\theta_j).$$

Moreover, according to (6.39) of [1], as $(\varepsilon_{i,j})$ has a moment of order > 2 , Lindeberg condition is satisfied for $\mathcal{M}_{n,j}(x)$. We can conclude from the central limit theorem for martingales given e.g. by Corollary 2.1.10 of [6] that for all $x \in [-1/2; 1/2]$ with $x \neq 0$,

$$(9.20) \quad \frac{\mathcal{M}_{n,j}(x)}{\sqrt{n^{1+\alpha}}} \xrightarrow{\mathcal{L}} \mathcal{N} \left(0, \frac{\sigma_j^2 \nu^2}{1+\alpha} (g(\theta_j + x) + g(\theta_j - x)) \right),$$

while, for $x = 0$,

$$(9.21) \quad \frac{\mathcal{M}_{n,j}(0)}{\sqrt{n^{1+\alpha}}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, 4 \frac{\sigma_j^2 \nu^2}{1+\alpha} g(\theta_j)\right).$$

Finally, it follows from (9.20) and (9.21) together with the previous convergence (9.6) and Slutsky's theorem that, for all $x \in [-1/2, 1/2]$ with $x \neq 0$,

$$(9.22) \quad \frac{1}{\sqrt{n^{1+\alpha}}} \sum_{j=1}^p \omega_j(x) \frac{\mathcal{M}_{n,j}(x)}{\mathcal{G}_{n,j}(x)} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\nu^2}{1+\alpha} \sum_{j=1}^p \frac{\sigma_j^2 \omega_j^2(x)}{a_j^2 (g(\theta_j + x) + g(\theta_j - x))}\right),$$

while, for $x = 0$,

$$(9.23) \quad \frac{1}{\sqrt{n^{1+\alpha}}} \sum_{j=1}^p \omega_j(0) \frac{\mathcal{M}_{n,j}(0)}{\mathcal{G}_{n,j}(0)} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\nu^2}{1+\alpha} \sum_{j=1}^p \frac{\sigma_j^2 \omega_j^2(0)}{a_j^2 g(\theta_j)}\right).$$

Then, the conjunction of (9.9), (9.12), (9.17) together with the two previous convergences (9.22) and (9.23) and Slutsky's theorem let us to conclude that, for all $x \in [-1/2, 1/2]$ with $x \neq 0$,

$$\sqrt{nh_n}(\hat{f}_n(x) - f(x)) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\nu^2}{1+\alpha} \sum_{j=1}^p \frac{\sigma_j^2 \omega_j^2(x)}{a_j^2 (g(\theta_j + x) + g(\theta_j - x))}\right),$$

while, for $x = 0$,

$$\sqrt{nh_n}(\hat{f}_n(0) - f(0)) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\nu^2}{1+\alpha} \sum_{j=1}^p \frac{\sigma_j^2 \omega_j^2(0)}{a_j^2 g(\theta_j)}\right),$$

which completes the proof of Theorem 6.2. \square

10. APPENDIX

In order to prove every identifiability conditions, let us consider that for a given vector of parameters (a, θ, v) satisfying (\mathcal{H}_4) and a given shape function f satisfying (\mathcal{H}_2) and (\mathcal{H}_3) , one can find another vector of parameters (a^*, θ^*, v^*) satisfying (\mathcal{H}_4) and an other shape function f^* satisfying (\mathcal{H}_2) and (\mathcal{H}_3) such that for all $1 \leq j \leq p$ and for all $x \in \mathbb{R}$, (2.1) is true, that is to say

$$(10.1) \quad a_j f(x - \theta_j) + v_j = a_j^* f^*(x - \theta_j^*) + v_j^*.$$

First of all, the periodicity of f and (\mathcal{H}_3) lead to $v_j = v_j^*$. Then, (10.1) becomes

$$(10.2) \quad a_j f(x - \theta_j) = a_j^* f^*(x - \theta_j^*).$$

For $j = 1$, the identifiability constraints $a_1 = a_1^*$ and $\theta_1 = \theta_1^*$ enables us to show that $f(x) = f^*(x)$. Then, (10.2) can be rewritten as

$$(10.3) \quad a_j f(x - \theta_j) = a_j^* f(x - \theta_j^*).$$

Hence, by denoting I_2 the integral of the square of f

$$I_2 = \int_0^1 f^2(x) dx,$$

and by squaring and integrating into (10.3), we obtain that

$$(10.4) \quad a_j^2 I_2 = (a_j^*)^2 I_2,$$

leading to

$$(10.5) \quad a_j^2 = (a_j^*)^2.$$

If $a_j = a_j^*$, then $f(x - \theta_j) = f(x - \theta_j^*)$ and it follows from the constraint

$$\max_{1 \leq j \leq p} |\theta_j| < 1/4$$

that $\theta_j = \theta_j^*$. Otherwise, if $a_j = -a_j^*$, then $f(x - \theta_j) + f(x - \theta_j^*) = 0$, which clearly leads to the identity $f(x) = f(x + 2(\theta_j - \theta_j^*))$. Therefore, the constraint

$$\max_{1 \leq j \leq p} |\theta_j| < 1/4$$

implies that $\theta_j = \theta_j^*$. Then, $f(x) = 0$, which is impossible. Finally, we have shown that

$$v = v^*, \quad \theta = \theta^*, \quad a = a^* \quad \text{and} \quad f = f^*,$$

leading to the identifiability of the model (1.1). This reasoning fits for the other sets of identifiability constraints. \square

Acknowledgements. *The author thanks Bernard Bercu for all his advices and for his thorough readings of the paper.*

REFERENCES

- [1] BERCU, B., AND FRAYSSE, P. A robbins-monro procedure for estimation in semiparametric regression models. *Ann. Statist.* 40 (2012).
- [2] CASTILLO, I., AND LOUBES, J.-M. Estimation of the distribution of random shifts deformation. *Math. Methods Statist.* 18 1 (2009), 21–42.
- [3] CHAABANE, F., AND MAAOUIA, F. Théorèmes limites avec poids pour les martingales vectorielles. *ESAIM PS* 4 (2000), 137–189.
- [4] DALALYAN, A. S., GOLUBEV, G. K., AND TSYBAKOV, A. B. Penalized maximum likelihood and semiparametric second-order efficiency. *Ann. Statist.* 34 1 (2006), 169–201.
- [5] DEVROYE, L., AND LUGOSI, G. *Combinatorial methods in density estimation*. Springer Series in Statistics. Springer-Verlag, New York, 2001.
- [6] DUFLO, M. *Random iterative models*, vol. 34 of *Applications of Mathematics*. Springer-Verlag, Berlin, 1997.
- [7] GAMBOA, F., LOUBES, J.-M., AND MAZA, E. Semi-parametric estimation of shifts. *Electron. J. Stat.* 1 (2007), 616–640.
- [8] GASSER, T., AND KNEIP, A. Searching for structure in curve sample. *Journal of the Am. Statist. Asso.*
- [9] HALL, P., AND HEYDE, C. C. *Martingale limit theory and its application*. Academic Press Inc. New York, 1980.
- [10] HÄRDLE, W., AND MARRON, J. S. Semiparametric comparison of regression curves. *Ann. Statist.* 18, 1 (1990), 63–89.
- [11] KNEIP, A., AND ENGEL, J. Model estimation in nonlinear regression under shape invariance. *Ann. Statist.* 23, 2 (1995), 551–570.
- [12] KUSHNER, H. J., AND YIN, G. G. *Stochastic approximation and recursive algorithms and applications*, vol. 35 of *Applications of Mathematics*. Springer-Verlag, New York, 2003.

- [13] LAWTON, W. H., SYLVESTRE, E. A., AND MAGGIO, M. S. Self modeling nonlinear regression. *Technometrics* 14 (1972), 513–532.
- [14] NADARAJA, È. On a regression estimate. *Teor. Verojatnost. i Primenen* 9 (1964), 157–159.
- [15] PELLETIER, M. On the almost sure asymptotic behaviour of stochastic algorithms. *Stochastic Process. Appl.* 78, 2 (1998), 217–244.
- [16] PELLETIER, M. Weak convergence rates for stochastic approximation with application to multiple targets and simulated annealing. *Annals of Appl. Proba.* 8, 1 (1998), 10–44.
- [17] RAMSAY, J. O., AND LI, X. Curve registration. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 60, 2 (1998), 351–363.
- [18] RAMSAY, J. O., AND SILVERMAN, B. W. *Functional data analysis*, second ed. Springer Series in Statistics. Springer, New York, 2005.
- [19] ROBBINS, H., AND MONRO, S. A stochastic approximation method. *Ann. Math. Statistics* 22 (1951), 400–407.
- [20] TRIGANO, T., ISSERLES, U., AND RITOV, Y. Semiparametric curve alignment and shift density estimation for biological data. *IEEE Trans. Signal Processing* 59 (2011), 1970–1984.
- [21] TSYBAKOV, A. B. *Introduction à l'estimation non-paramétrique*, vol. 41 of *Mathématiques & Applications (Berlin)*. Springer-Verlag, Berlin, 2004.
- [22] VIMOND, M. Efficient estimation for a subclass of shape invariant models. *Ann. Statist.* 38, 3 (2010), 1885–1912.
- [23] WANG, Y., KE, C., AND BROWN, M. B. Shape-invariant modeling of circadian rhythms with random effects and smoothing spline ANOVA decompositions. *Biometrics* 59, 4 (2003), 804–812.
- [24] WATSON, G. Smooth regression analysis. *Sankhya Ser. A*.
E-mail address: philippe.fraysse@math.u-bordeaux1.fr

UNIVERSITÉ BORDEAUX 1, INSTITUT DE MATHÉMATIQUES DE BORDEAUX, UMR CNRS 5251, AND INRIA BORDEAUX, TEAM ALEA, 351 COURS DE LA LIBÉRATION, 33405 TALENCE CEDEX, FRANCE.